

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Chiron Corporation
- (ii) TITLE OF INVENTION: Methods for Administration of Recombinant Gene Delivery Vehicles for Treatment of Hemophilia and Other Disorders
- (iii) NUMBER OF SEQUENCES: 83
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Chiron Corporation
 - (B) STREET: 4560 Horton Street
 - (C) CITY: Emeryville
 - (D) STATE: California
 - (E) COUNTRY: U.S.A.
 - (F) ZIP: 94608
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Kruse, Norman J.
 - (B) REGISTRATION NUMBER: 35,235
 - (C) REFERENCE/DOCKET NUMBER: 1155.005
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (510) 923-3520
 - (B) TELEFAX: (510) 655-3542

- (2) INFORMATION FOR SEQ ID NO:1:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
GAGAGATGGG GGAGGCTAAC TGAG 24
- (2) INFORMATION FOR SEQ ID NO:2:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 28 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:
GATCCTCAGT TAGCCTCCCC CATCTCTC 28
- (2) INFORMATION FOR SEQ ID NO:3:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 35 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:
TATATCTCCA GATGAGGTAC ATGATTTTAG GCTTG 35
- (2) INFORMATION FOR SEQ ID NO:4:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 40 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:
TATATATCGA TTCAAGGCAT TTTCTTTTCA TCAATAAAAC 40
- (2) INFORMATION FOR SEQ ID NO:5:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 37 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:
TCGAGGATCC GCCC GGCGG CCGCATCGAT GTCGACG 37
- (2) INFORMATION FOR SEQ ID NO:6:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 35 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:
CGCGTCGACA TCGATGCGGC CGCCCGGGCG GATCC 35

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 77 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AGTGAATTCG AGCTCGGTAC CCGGGGATCC TCTAGAGTCG ACCTGCAGGC ATGCAAGCTT
GGCGTAATCA TGGTCAT

60

77

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Ala Arg Glu Met Gly Glu Ala Asn
1 5

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CCCGAGAGAT GGGGGAGGCT AACTGAG

27

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GGGCTCTCTA CCCCTCCGA TTGACACCTA G

31

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Thr Ile Met Thr Met
1 5

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CCCTGTGCCT TATTTGAACT AACC

24

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CCCACCACAA CCACATATCC CTCC

24

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CCAGTCCTCC GATTGACTG

19

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8332 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GCGCCAGTCC TCCGATTGAC TGAGTCGCCC GGGTACCCGT GTATCCAATA AACCTCTTG 60

CAGTTGCATC CGACTTGTGG TCTCGCTGTT CCTTGGGAGG GTCTCCTCTG AGTGATTGAC 120

TACCCGTCAG CGGGGGTCTT TCATTTGGGG GCTCGTCCGG GATCGGGAGA CCCCTGCCCA 180

GGGACCACCG ACCCACCACC GGGAGGTAAG CTGGCCAGCA ACTTATCTGT GTCTGTCCGA 240

TTGTCTAGTG TCTATGACTG ATTTTATGCG CCTGCGTCGG TACTAGTTAG CTAAGTAGCT 300

CTGTATCTGG CGGACCCGTG GTGGAAGTGA CGAGTTCGGA ACACCCGGCC GCAACCCTGG 360

GAGACGTCCC AGGGACTTCG GGGGCCGTTT TTGTGGCCCG ACCTGAGTCC AAAAATCCCG 420

ATCGTTTTGG ACTCTTTGGT GCACCCCCCT TAGAGGAGGG ATATGTGGTT CTGGTAGGAG 480

ACGAGAACCT AAAACAGTTC CCGCCTCCGT CTGAATTTTT GCTTTCGGTT TGGGACCGAA 540

GCCGCGCCGC GCGTCTTGTC TGCTGCAGCA TCGTTCTGTG TTGTCTCTGT CTGACTGTGT 600

TTCTGTATTT GTCTGAGAAT ATGGGCCAGA CTGTTACCAC TCCCTTAAGT TTGACCTTAG 660

GCTACTGGAA	AGATGTCGAG	CGGATCGCTC	ACAACCAGTC	GGTAGATGTC	AAGAAGAGAC	720
GTTGGGTTAC	CTTCTGCTCT	GCAGAATGGC	CAACCTTTAA	CGTCGGATGG	CCGCGAGACG	780
GCACCTTTAA	CCGAGACCTC	ATCACCCAGG	TTAAGATCAA	GGTCTTTTCA	CCTGGCCCCG	840
ATGGACACCC	AGACCAGGTC	CCCTACATCG	TGACCTGGGA	AGCCTTGGCT	TTTGACCCCC	900
CTCCCTGGGT	CAAGCCCTTT	GTACACCCTA	AGCCTCCGCC	TCCTCTTCCT	CCATCCGCCC	960
CGTCTCTCCC	CCTTGAACCT	CCTCGTTCGA	CCCCGCCTCG	ATCCTCCCTT	TATCCAGCCC	1020
TCACTCCTTC	TCTAGGCGCC	AAACCTAAAC	CTCAAGTTCT	TTCTGACAGT	GGGGGGCCGC	1080
TCATCGACCT	ACTTACAGAA	GACCCCCCGC	CTTATAGGGA	CCCAAGACCA	CCCCCTTCCG	1140
ACAGGGACGG	AAATGGTGGA	GAAGCGACCC	CTGCGGGAGA	GGCACC GGAC	CCCTCCCCAA	1200
TGGCATCTCG	CCTACGTGGG	AGACGGGAGC	CCCCTGTGGC	CGACTCCACT	ACCTCGCAGG	1260
CATTCCCCCT	CCGCGCAGGA	GGAAACGGAC	AGCTTCAATA	CTGGCCGTTT	TCCTCTTCTG	1320
ACCTTTACAA	CTGGAAAAAT	AATAACCCTT	CTTTTTCTGA	AGATCCAGGT	AAACTGACAG	1380
CTCTGATCGA	GTCTGTTCTC	ATCACCCATC	AGCCACCTG	GGACGACTGT	CAGCAGCTGT	1440
TGGGGACTCT	GCTGACCGGA	GAAGAAAAAC	AACGGGTGCT	CTTAGAGGCT	AGAAAGGCGG	1500
TGCGGGGCGA	TGATGGGCGC	CCCACTCAAC	TGCCCAATGA	AGTCGATGCC	GCTTTTCCCC	1560
TCGAGCGCCC	AGACTGGGAT	TACACCACCC	AGGCAGGTAG	GAACCACCTA	GTCCACTATC	1620
GCCAGTTGCT	CCTAGCGGGT	CTCCAAAACG	CGGGCAGAAG	CCCCACCAAT	TTGGCCAAGG	1680
TAAAAGGAAT	AACACAAGGG	CQCAATGAGT	CTCCCTCGGC	CTTCCTAGAG	AGACTTAAGG	1740
AAGCCTATCG	CAGGTACACT	CCTTATGACC	CTGAGGACCC	AGGGCAAGAA	ACTAATGTGT	1800
CTATGTCTTT	CATTTGGCAG	TCTGCCCCAG	ACATTGGGAG	AAAGTTAGAG	AGGTTAGAAG	1860
ATTTAAAAAA	CAAGACGCTT	GGAGATTTGG	TTAGAGAGGC	AGAAAAGATC	TTTAATAAAC	1920
GAGAAACCCC	GGAAGAAAGA	GAGGAACGTA	TCAGGAGAGA	AACAGAGGAA	AAAGAAGAAC	1980
GCCGTAGGAC	AGAGGATGAG	CAGAAAGAGA	AAGAAAGAGA	TCGTAGGAGA	CATAGAGAGA	2040
TGAGCAAGCT	ATTGGCCACT	GTCGTTAGTG	GACAGAAACA	GGATAGACAG	GGAGGAGAAC	2100
GAAGGAGGTC	CCAACCTGAT	CGCGACCAGT	GTGCCTACTG	CAAAGAAAAG	GGGCACTGGG	2160
CTAAAGATTG	TCCCAAGAAA	CCACGAGGAC	CTCGGGGACC	AAGACCCAG	ACCTCCCTCC	2220
TGACCCTAGA	TGACTAGGGA	GGTCAGGGTC	AGGAGCCCCC	CCCTGAACCC	AGGATAACCC	2280
TCAAAGTCGG	GGGGCAACCC	GTCACCTTCC	TGGTAGATAC	TGGGGCCCAA	CACTCCGTGC	2340
TGACCCAAAA	TCCTGGACCC	CTAAGTGATA	AGTCTGCCTG	GGTCCAAGGG	GCTACTGGAG	2400
GAAAGCGGTA	TCGCTGGACC	ACGGATCGCA	AAGTACATCT	AGCTACCGGT	AAGGTCACCC	2460
ACTCTTTCCT	CCATGTACCA	GACTGTCCCT	ATCCTCTGTT	AGGAAGAGAT	TTGCTGACTA	2520

AACTAAAAGC	CCAAATCCAC	TTTGAGGGAT	CAGGAGCTCA	GGTTATGGGA	CCAATGGGGC	2580
AGCCCCTGCA	AGTGTTGACC	CTAAATATAG	AAGATGAGCA	TCGGCTACAT	GAGACCTCAA	2640
AAGAGCCAGA	TGTTTCTCTA	GGGTCCACAT	GGCTGTCTGA	TTTTCCTCAG	GCCTGGGCGG	2700
AAACCGGGGG	CATGGGACTG	GCAGTTCGCC	AAGCTCCTCT	GATCATACTT	CTGAAAGCAA	2760
CCTCTACCCC	CGTGTCCATA	AAACAATACC	CCATGTCACA	AGAAGCCAGA	CTGGGGATCA	2820
AGCCCCACAT	ACAGAGACTG	TTGGACCAGG	GAATACTGGT	ACCCTGCCAG	TCCCCCTGGA	2880
ACACGCCCCCT	GCTACCCGTT	AAGAAACCAG	GGACTAATGA	TTATAGGCCT	GTCCAGGATC	2940
TGAGAGAAGT	CAACAAGCGG	GTGGAAGACA	TCCACCCAC	CGTGCCCAAC	CCTTACAACC	3000
TCTTGAGCGG	GCTCCCACCG	TCCCACCAGT	GGTACACTGT	GCTTGATTTA	AAGGATGCCT	3060
TTTTCCTGCCT	GAGACTCCAC	CCCACCAGTC	AGCCTCTCTT	CGCCTTTGAG	TGGAGAGATC	3120
CAGAGATGGG	AATCTCAGGA	CAATTGACCT	GGACCAGACT	CCCACAGGGT	TTCAAAAACA	3180
GTCCCACCCT	GTTTGATGAG	GCACTGCACA	GAGACCTAGC	AGACTTCCGG	ATCCAGCACC	3240
CAGACTTGAT	CCTGCTACAG	TACGTGGATG	ACTTACTGCT	GGCCGCCACT	TCTGAGCTAG	3300
ACTGCCAACA	AGGTACTCGG	GCCCTGTTAC	AAACCCTAGG	GAACCTCGGG	TATCGGGCCT	3360
CGGCCAAGAA	AGCCCAAATT	TGCCAGAAAC	AGGTCAAGTA	TCTGGGGTAT	CTTCTAAAAG	3420
AGGGTCAGAG	ATGGCTGACT	GAGGCCAGAA	AAGAGACTGT	GATGGGGCAG	CCTACTCCGA	3480
AGACCCCTCG	ACAACCTAAG	GAGTCTCTAG	GGACGGCAGG	CTTCTGTCTG	CTCTGGATCC	3540
CTGGGTTTGC	AGAAATGGCA	GCCCCCTTGT	ACCTCTCAC	CAAAACGGGG	ACTCTGTTTA	3600
ATTGGGGCCC	AGACCAACAA	AAGGCTATC	AAGAAATCAA	GCAAGCTCTT	CTAACTGCCC	3660
CAGCCCTGGG	GTGCCAGAT	TTGACTAAGC	CCTTTGAACT	CTTTGTCTGAC	GAGAAGCAGG	3720
GCTACGCCAA	AGGTGTCTTA	ACGCAAAAAC	TGGGACCTTG	GCGTCGGCCG	GTGGCCTACC	3780
TGTCCAAAAA	GCTAGACCCA	GTAGCAGCTG	GGTGGCCCCC	TTGCCTACGG	ATGGTAGCAG	3840
CCATTGCCGT	ACTGACAAAG	GATGCAGGCA	AGCTAACCAT	GGGACAGCCA	CTAGTCATTC	3900
TGGCCCCCCA	TGCAGTAGAG	GCACTAGTCA	AACAACCCCC	CGACCGCTGG	CTTTCCAACG	3960
CCCGGATGAC	TCACTATCAG	GCCTTGCTTT	TGGACACGGA	CCGGGTCCAG	TTCGGACCGG	4020
TGGTAGCCCT	GAACCCGGCT	ACGCTGCTCC	CACTGCCTGA	GGAAGGGCTG	CAACACAACCT	4080
GCCTTGATAT	CCTGGCCGAA	GCCCACGGAA	CCCGACCCGA	CCTAACGGAC	CAGCCGCTCC	4140
CAGACGCCGA	CCACACCTGG	TACACGGATG	GAAGCAGTCT	CTTACAAGAG	GGACAGCGTA	4200
AGGCGGGAGC	TGCGGTGACC	ACCGAGACCG	AGGTAATCTG	GGCTAAAGCC	CTGCCAGCCG	4260
GGACATCCGC	TCAGCGGGCT	GAAGTGATAG	CACTCACCCA	GGCCCTAAAG	ATGGCAGAAG	4320

GTAAGAAGCT	AAATGTTTAT	ACTGATAGCC	GTTATGCTTT	TGCTACTGCC	CATATCCATG	4380
GAGAAATATA	CAGAAGGCGT	GGGTTGCTCA	CATCAGAAGG	CAAAGAGATC	AAAAATAAAG	4440
ACGAGATCTT	GGCCCTACTA	AAAGCCCTCT	TTCTGCCCAA	AAGACTTAGC	ATAATCCATT	4500
GTCCAGGACA	TCAAAAGGGA	CACAGCGCCG	AGGCTAGAGG	CAACCGGATG	GCTGACCAAG	4560
CGGCCCCAAA	GGCAGCCATC	ACAGAGACTC	CAGACACCTC	TACCCTCCTC	ATAGAAAATT	4620
CATCACCTTA	CACCTCAGAA	CATTTTCATT	ACACAGTGAC	TGATATAAAG	GACCTAACCA	4680
AGTTGGGGGC	CATTTATGAT	AAAACAAAGA	AGTATTGGGT	CTACCAAGGA	AAACCTGTGA	4740
TGCCTGACCA	GTTTACTTTT	GAATTATTAG	ACTTTCTTCA	TCAGCTGACT	CACCTCAGCT	4800
TCTCAAAAAT	GAAGGCTCTC	CTAGAGAGAA	GCCACAGTCC	CTACTACATG	CTGAACCGGG	4860
ATCGAACACT	CAAAAATATC	ACTGAGACCT	GCAAAGCTTG	TGCACAAGTC	AACGCCAGCA	4920
AGTCTGCCGT	TAAACAGGGA	ACTAGGGTCC	GCGGGCATCG	GCCCGGCACT	CATTGGGAGA	4980
TCGATTTTAC	CGAGATAAAG	CCCGGATTGT	ATGGCTATAA	ATATCTTCTA	GTTTTTATAG	5040
ATACCTTTTC	TGGCTGGATA	GAAGCCTTCC	CAACCAAGAA	AGAAACCGCC	AAGGTCGTAA	5100
CCAAGAAGCT	ACTAGAGGAG	ATCTTCCCCA	GGTTCGGCAT	GCCTCAGGTA	TTGGGAACTG	5160
ACAATGGGCC	TGCCTTCGTC	TCCAAGGTGA	GTCAGACAGT	GGCCGATCTG	TTGGGGATTG	5220
ATTGGAAATT	ACATTGTGCA	TACAGACCCC	AAAGCTCAGG	CCAGGTAGAA	AGAATGAATA	5280
GAACCATCAA	GGAGACTTTA	ACTAAATTAA	CGCTTGCAAC	TGGCTCTAGA	GACTGGGTGC	5340
TCCTACTCCC	CTTAGCCCTG	TACCGAGCCC	GCAACACGCC	GGGCCCCCAT	GGCCTCACCC	5400
CATATGAGAT	CTTATATGGG	GCACCCCCGC	CCCTTGTAAG	CTTCCCTGAC	CCTGACATGA	5460
CAAGAGTTAC	TAACAGCCCC	TCTCTCCAAG	CTCACTTACA	GGCTCTCTAC	TTAGTCCAGC	5520
ACGAAGTCTG	GAGACCTCTG	GCGGCAGCCT	ACCAAGAACA	ACTGGACCGA	CCGGTGGTAC	5580
CTCACCTTA	CCGAGTCGGC	GACACAGTGT	GGGTCCGCCG	ACACCAGACT	AAGAACCTAG	5640
AACCTCGCTG	GAAAGGACCT	TACACAGTCC	TGCTGACCAC	CCCCACCGCC	CTCAAAGTAG	5700
ACGGCATCGC	AGCTTGATA	CACGCCGCCC	ACGTGAAGGC	TGCCGACCCC	GGGGGTGGAC	5760
CATCCTCTAG	ACTGACATGG	CGCGTTCAAC	GCTCTCAAAA	CCCCTTAAAA	ATAAGGTTAA	5820
CCCGCGAGGC	CCCCTAATCC	CCTTAATTCT	TCTGATGCTC	AGAGGGGTCA	GTAAGGTTTC	5880
GCCCGGCTCC	AGTCCTCATC	AAGTCTATAA	TATCACCTGG	GAGGTAACCA	ATGGAGATCG	5940
GGAGACGGTA	TGGGCAACTT	CTGGCAACCA	CCCTCTGTGG	ACCTGGTGGC	CTGACCTTAC	6000
CCCAGATTTA	TGTATGTTAG	CCCACCATGG	ACCATCTTAT	TGGGGGCTAG	AATATCAATC	6060
CCCTTTTTCT	TCTCCCCCGG	GGCCCCCTTG	TTGCTCAGGG	GGCAGCAGCC	CAGGCTGTTC	6120
CAGAGACTGC	GAAGAACCTT	TAACCTCCCT	CACCCCTCGG	TGCAAACTG	CCTGGAACAG	6180

ACTCAAGCTA	GACCAGACAA	CTCATAAATC	AAATGAGGGA	TTTTATGTTT	GCCCCGGGCC	6240
CCACCGCCCC	CGAGAATCCA	AGTCATGTGG	GGGTCCAGAC	TCCTTCTACT	GTGCCTATTG	6300
GGGCTGTGAG	ACAACCGGTA	GAGCTTACTG	GAAGCCCTCC	TCATCATGGG	ATTTTCATCAC	6360
AGTAAACAAC	AATCTCACCT	CTGACCAGGC	TGTCCAGGTA	TGCAAAGATA	ATAAGTGGTG	6420
CAACCCCTTA	GTTATTCGGT	TTACAGACGC	CGGGAGACGG	GTTACTTCCT	GGACCACAGG	6480
ACATTACTGG	GGCTTACGTT	TGTATGTCTC	CGGACAAGAT	CCAGGGCTTA	CATTTGGGAT	6540
CCGACTCAGA	TACCAAAATC	TAGGACCCCG	CGTCCCAATA	GGGCCAAACC	CCGTTCTGGC	6600
AGACCAACAG	CCACTCTCCA	AGCCCAAACC	TGTTAAGTCG	CCTTCAGTCA	CCAAACCACC	6660
CAGTGGGACT	CCTCTCTCCC	CTACCCAACT	TCCACCGGCG	GGAACGGAAA	ATAGGCTGCT	6720
AAACTTAGTA	GACGGAGCCT	ACCAAGCCCT	CAACCTCACC	AGTCCTGACA	AAACCCAAGA	6780
GTGCTGGTTG	TGTCTAGTAG	CGGGACCCCC	CTACTACGAA	GGGGTTGCCG	TCCTGGGTAC	6840
CTACTCCAAC	CATACCTCTG	CTCCAGCCAA	CTGCTCCGTG	GCCTCCCAAC	ACAAGTTGAC	6900
CCTGTCCGAA	GTGACCGGAC	AGGGACTCTG	CATAGGAGCA	GTTCCCAAAA	CACATCAGGC	6960
CCTATGTAAT	ACCACCCAGA	CAAGCAGTCG	AGGGTCCTAT	TATCTAGTTG	CCCCTACAGG	7020
TACCATGTGG	GCTTGTAAGT	CCGGGCTTAC	TCCATGCATC	TCCACCACCA	TACTGAACCT	7080
TACCACTGAT	TATTGTGTTC	TTGTGCAACT	CTGGCCAAGA	GTCACCTATC	ATTCCCCCAG	7140
CTATGTTTAC	GGCCTGTTTG	AGAGATCCAA	CCGACACAAA	AGAGAACCGG	TGTCGTTAAC	7200
CCTGGCCCTA	TTATTGGGTG	GACTAACCAT	GGGGGAATT	GCCGCTGGAA	TAGGAACAGG	7260
GACTACTGCT	CTAATGGCCA	CTCAGCAATT	CCAGCAGCTC	CAAGCCGCAG	TACAGGATGA	7320
TCTCAGGGAG	GTTGAAAAAT	CAATCTCTAA	CCTAGAAAAG	TCTCTCACTT	CCCTGTCTGA	7380
AGTTGTCTTA	CAGAATCGAA	GGGGCCTAGA	CTTGTTATTT	CTAAAAGAAG	GAGGGCTGTG	7440
TGCTGCTCTA	AAAGAAGAAT	GTTGCTTCTA	TGCGGACCAC	ACAGGACTAG	TGAGAGACAG	7500
CATGGCCAAA	TTGAGAGAGA	GGCTTAATCA	GAGACAGAAA	CTGTTTGAGT	CAACTCAAGG	7560
ATGGTTTGAG	GGACTGTTTA	ACAGATCCCC	TTGGTTTACC	ACCTTGATAT	CTACCATTAT	7620
GGGACCCCTC	ATTGTACTCC	TAATGATTTT	GCTCTTCGGA	CCCTGCATTC	TTAATCGATT	7680
AGTCCAATTT	GTTAAAGACA	GGATATCAGT	GGTCCAGGCT	CTAGTTTTGA	CTCAACAATA	7740
TCACCAGCTG	AAGCCTATAG	AGTACGAGCC	ATAGATAAAA	TAAAAGATTT	TATTTAGTCT	7800
CCAGAAAAAG	GGGGGAATGA	AAGACCCAC	CTGTAGGTTT	GGCAAGCTAG	CTTAAGTAAC	7860
GCCATTTTGC	AAGGCATGGA	AAAATACATA	ACTGAGAATA	GAGAAGTTCA	GATCAAGGTC	7920
AGGAACAGAT	GGAACAGCTG	AATATGGGCC	AAACAGGATA	TCTGTGGTAA	GCAGTTCCTG	7980

CCCCGGCTCA GGGCCAAGAA CAGATGGAAC AGCTGAATAT GGGCCAAACA GGATATCTGT 8040
GGTAAGCAGT TCCTGCCCCG GCTCAGGGCC AAGAACAGAT GGTCCCCAGA TGCGGTCCAG 8100
CCCTCAGCAG TTTCTAGAGA ACCATCAGAT GTTTCAGGG TGCCCCAAGG ACCTGAAATG 8160
ACCCTGTGCC TTATTTGAAC TAACCAATCA GTTCGCTTCT CGCTTCTGTT CGCGCGCTTC 8220
TGCTCCCCGA GCTCAATAAA AGAGCCCACA ACCCCTCACT CGGGGCGCCA GTCCTCCGAT 8280
TGA CTGAGTC GCCCGGGTAC CCGTGTATCC AATAAACCCCT CTTGCAGTTG CA 8332

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:
GGGAGTGGTA ACAGTCTGGC CTTAATTCTC AG

32

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:
CGGTCGACCT CGAGAATTAA TTC

23

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:
CTGGGAGACG TCCCAGGGAC TTC

23

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:
GGCCAGACTG TTACCACTCC CTGAAGTTTG AC

32

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:
CATCGATAAA ATAAAAGATT TTATTTAGTC

30

- (2) INFORMATION FOR SEQ ID NO:21:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:
CAAATGAAAG ACCCCCGCTG AC 22
- (2) INFORMATION FOR SEQ ID NO:22:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 39 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:
GAAGCTTCTC CCAGAACCCA CCAGTCTTGA AACGCCATC 39
- (2) INFORMATION FOR SEQ ID NO:23:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:
GTACCAGCTT TTGGTCTCAT CAAAG 25
- (2) INFORMATION FOR SEQ ID NO:24:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 36 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:
CTCCTCGAGC TAAAGATATT TTAGAGAAGA ATTAAC 36
- (2) INFORMATION FOR SEQ ID NO:25:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:
TTCCTCTGGA CAGCTGTCTA CTTTG 25
- (2) INFORMATION FOR SEQ ID NO:26:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 51 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:
TCGAGGGGCC CAGATCTGCG GCCGCTCGCG AGTCGACAAG CTTGGATCCA T 51

- (2) INFORMATION FOR SEQ ID NO:27:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 49 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:
CGATGGATCC AAGCTTGTCG ACTCGCGAGC GGCCGCAGAT CTGGGCCCC 49
- (2) INFORMATION FOR SEQ ID NO:28:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:
CACCGTCGTC GACTTATGCT 20
- (2) INFORMATION FOR SEQ ID NO:29:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 34 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:
GACCGTCGAC TCAATTCTGG GAGAAGCTTC TTGG 34
- (2) INFORMATION FOR SEQ ID NO:30:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:
CACCGTCGTC GACTTATGCT 20
- (2) INFORMATION FOR SEQ ID NO:31:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:
CAACGCTCGA GAAGCAGAAT CGCAAAAGGC 30
- (2) INFORMATION FOR SEQ ID NO:32:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 32 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:
TCGGCTCGAG GCATCAACGG GAAATAACTC GT 32

- (2) INFORMATION FOR SEQ ID NO:33:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 31 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:
CCGACTCGAG TCAGTAGAGG TCCTGTGCCT C 31
- (2) INFORMATION FOR SEQ ID NO:34:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 26 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:
GCGACTCGAG CATGGGGCCC TGGGGC 26
- (2) INFORMATION FOR SEQ ID NO:35:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 21 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:
GCACTGGAAT TCGTCAGGGC G 21
- (2) INFORMATION FOR SEQ ID NO:36:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:
CGCGCCGCTC GAGTCTACAA TGGCCTTGAC CTTTGCTTTA CTGG 44
- (2) INFORMATION FOR SEQ ID NO:37:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:
GCGCCCATCG ATTCATTCCT TACTTCTTAA ACTTTCTTGC AAG 43
- (2) INFORMATION FOR SEQ ID NO:38:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 47 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:
CGCGCCGCTC GAGCATCCAA TGGCCCTGTC CTTTCTTTA CTTATGG 47

- (2) INFORMATION FOR SEQ ID NO:39:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 39 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:
 CCATCGATTC AATCCTTCCT CCTTAATCTT TTTTGCAAG 39
- (2) INFORMATION FOR SEQ ID NO:40:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 39 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:
 CCGGATCCTC TACAATGGCC TTGACCTTG CTTTACTGG 39
- (2) INFORMATION FOR SEQ ID NO:41:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 46 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:
 CGCGCCGGCG GCCGCTCATT CCTTACTTCT TAACTTTCT TGCAAG 46
- (2) INFORMATION FOR SEQ ID NO:42:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 42 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:
 CCGGATCCCA TCCAATGGCC CTGTCCTTTT CTTTACTTAT GG 42
- (2) INFORMATION FOR SEQ ID NO:43:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 46 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:
 CGCGCCGGCG GCCGCTCAAT CCTTCCTCCT TAATCTTTTT TGCAAG 46
- (2) INFORMATION FOR SEQ ID NO:44:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 9080 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:
 GGGGGGGGGG GGGGGGGGGG GGGTGAGCAC ATCCAGTGGG TAAAGTTCCT TAAAATGCTC 60
 TGCAAAGAAA TTGGGACTTT TCATTAAATC AGAAATTTTA CTTTTTTCCC CTCCTGGGAG 120
 CTAAAGATAT TTTAGAGAAG AATTAACCTT TTGCTTCTCC AGTTGAACAT TTGTAGCAAT 180

AAGTCATGCA	AATAGAGCTC	TCCACCTGCT	TCTTTCTGTG	CCTTTTGCGA	TTCTGCTTTA	240
GTGCCACCAG	AAGATACTAC	CTGGGTGCAG	TGGAACCTGT	ATGGGACTAT	ATGCAAAGTG	300
ATCTCGGTGA	GCTGCCCTGT	GACGCAAGAT	TTCCCTCCTAG	AGTGCCAAAA	TCTTTTCCAT	360
TCAACACCTC	AGTCGTGTAC	AAAAAGACTC	TGTTTGTAGA	ATTCACGGAT	CACCTTTTCA	420
ACATCGCTAA	GCCAAGGCCA	CCCTGGATGG	GTCTGCTAGG	TCCTACCATC	CAGGCTGAGG	480
TTTATGATAC	AGTGGTCATT	ACACTTAAGA	ACATGGCTTC	CCATCCTGTC	AGTCTTCATG	540
CTGTTGGTGT	ATCCTACTGG	AAAGCTTCTG	AGGGAGCTGA	ATATGATGAT	CAGACCAGTC	600
AAAGGGAGAA	AGAAGATGAT	AAAGTCTTCC	CTGGTGGAAG	CCATACATAT	GTCTGGCAGG	660
TCCTGAAAAG	GAATGGTCCA	ATGGCCTCTG	ACCCACTGTG	CCTTACCTAC	TCATATCTTT	720
CTCATGTGGA	CCTGGTAAAA	GACTTGAATT	CAGGCCTCAT	TGGAGCCCTA	CTAGTATGTA	780
GAGAAGGGAG	TCTGGCCAAG	GAAAAGACAC	AGACCTTGCA	CAAAATTTATA	CTACTTTTGT	840
CTGTATTTGA	TGAAGGGAAA	AGTTGGCACT	CAGAAAACAAA	GAACTCCTTG	ATGCAGGATA	900
GGGATGCTGC	ATCTGCTCGG	GCCTGGCCTA	AAATGCACAC	AGTCAATGGT	TATGTAAACA	960
GGTCTCTGCC	AGGTCTGATT	GGATGCCACA	GGAAATCAGT	CTATTGGCAT	GTGATTGGAA	1020
TGGGCACCAC	TCCTGAAGTG	CACTCAATAT	TCCTCGAAGG	TCACACATTT	CTTGTGAGGA	1080
ACCATCGCCA	GGCGTCCTTG	GAAATCTCGC	CAATAACTTT	CCTTACTGCT	CAAACACTCT	1140
TGATGGACCT	TGGACAGTTT	CTACTGTTTT	GTCATATCTC	TTCCCACCAA	CATGATGGCA	1200
TGGAAGCTTA	TGTCAAAGTA	GACAGCTGTC	CAGAGGAACC	CCAACACGA	ATGAAAAATA	1260
ATGAAGAAGC	GTGAAGACTAT	GATGATGATC	TTACTGATTC	TGAAATGGAT	GTGGTCAGGT	1320
TTGATGATGA	CAACTCTCCT	TCCTTTTATCC	AAATTGCTC	AGTTGCCAAG	AAGCATCCTA	1380
AACTTTGGGT	ACATTACATT	GCTGCTGAAG	AGGAGGACTG	GGACTATGCT	CCCTTAGTCC	1440
TCGCCCCCGA	TGACAGAAGT	TATAAAAGTC	AATATTTGAA	CAATGGCCCT	CAGCGGATTG	1500
GTAGGAAGTA	CAAAAAAGTC	CGATTTTATGG	CATACACAGA	TGAAACCTTT	AAGACTCGTG	1560
AAGCTATTCA	GCATGAATCA	GGAACTCTTG	GACCTTTACT	TTATGGGGAA	GTTGGAGACA	1620
CACTGTTGAT	TATATTTAAG	AATCAAGCAA	GCAGACCATA	TAACATCTAC	CCTCACGGAA	1680
TCACTGATGT	CCGTCTTTTG	TATTTCAAGGA	GATTACCAAA	AGGTGTAAAA	CATTTGAAGG	1740
ATTTTCCAAT	TCTGCCAGGA	GAAATATTCA	AATATAAATG	GACAGTGACT	GTAGAAGATG	1800
GGCCAACATA	ATCAGATCCT	CGGTGCCTGA	CCCGCTATTA	CTCTAGTTTC	GTTAATATGG	1860
AGAGAGATCT	AGCTTCAGGA	CTCATTGGCC	CTCTCCTCAT	CTGCTACAAA	GAATCTGTAG	1920
ATCAAAGAGG	AAACCAGATA	ATGTCAGACA	AGAGGAATGT	CATCCTGTTT	TCTGTATTTG	1980
ATGAGAACCG	AAGCTGGTAC	CTCACAGAGA	ATATACAACG	CTTTCTCCCC	AATCCAGCTG	2040
GAGTGCAGCT	TGAGGATCCA	GAGTTCCAAG	CCTCCAACAT	CATGCACAGC	ATCAATGGCT	2100
ATGTTTTTGA	TAGTTTGCAG	TTGTCAAGTT	GTTTGCATGA	GGTGGCATAC	TGGTACATTC	2160
TAAGCATTTG	AGCACAGACT	GACTTCCCTT	CTGTCTTCTT	CTCTGGATAT	ACCTTCAAAC	2220
ACAAAATGGT	CTATGAAGAC	ACACTCACCC	TATTTCCATT	CTCAGGAGAA	ACTGTCTTCA	2280
TGTCGATGGA	AAACCCAGGT	CTATGGATTC	TGGGGTGCCA	CAACTCAGAC	TTTCGGAACA	2340
GAGGCATGAC	CGCCTTACTG	AAGGTTTCTA	GTTGTGACAA	GAACACTGGT	GATTATTACG	2400
AGGACAGTTA	TGAAGATATT	TCAGCATACT	TGCTGAGTAA	AAACAATGCC	ATTGAACCAA	2460
GAAGCTTCTC	CCAGAATTCA	AGACACCCTA	GCCTAGGACA	AAAGCAATTT	AATGCCACCA	2520
CAATTTCCAGA	AAATGACATA	GAGAAGACTG	ACCCTTGGTT	TGCACACAGA	ACACCTATGC	2580
CTAAAATACA	AAATGTCTCC	TCTAGTGATT	TGTTGATGCT	CTTGCGACAG	AGTCCTACTC	2640
CACATGGGCT	ATCCTTATCT	GATCTCCAAG	AAGCCAAATA	TGAGACTTTT	TCTGATGATC	2700
CATCACCTGG	AGCAATAGAC	AGTAATAACA	GCCTGTCTGA	AATGACACAC	TTCAGGCCAC	2760
AGCTCCATCA	CAGTGGGGAC	ATGGTATTTA	CCCCTGAGTC	AGGCCTCCAA	TTAAGATTAA	2820
ATGAGAAACT	GGGGACAAC	GCAGCAACAG	AGTTGAAGAA	ACTTGATTTT	AAAGTTTCTA	2880
GTACATCAAA	TAATCTGATT	TCAACAAATC	CATCAGACAA	TTTGGCAGCA	GGTACTGATA	2940
ATACAAGTTC	CTTAGGACCC	CCAAGTATGC	CAGTTCATTA	TGATAGTCAA	TGATATACCA	3000
CTCTATTTGG	CAAAAAGTCA	TCTCCCTTAA	CTGAGTCTGG	TGGACCTCTG	AGCTTGAGTG	3060
AAGAAAATAA	TGATTCAAAG	TTGTTAGAAT	CAGGTTTAAT	GAATAGCCAA	GAAAGTTTCT	3120
GGGGAAAAAA	TGTATCGTCA	ACAGAGAGTG	GTAGGTTATT	TAAAGGGAAA	AGAGCTCATG	3180
GACCTGCTTT	GTTGACTAAA	GATAATGCCT	TATTTCAAAG	TAGCATCTCT	TTGTTAAAGA	3240
CAAACAAAAC	TTCCAATAAT	TCAGCAACTA	ATAGAAAAGAC	TCACATTGAT	GGCCCATCAT	3300
TATTAATTGA	GAATAGTCCA	TCAGTCTGGC	AAAAATATAT	AGAAAGTGAC	ACTGAGTTTA	3360
AAAAAGTGAC	ACCTTTGATT	CATGACAGAA	TGCTTATGGA	CAAAAATGCT	ACAGCTTTGA	3420
GGCTAAATCA	AAAACTACTT	TATGTCAAAAT	CATCAAAAAA	CATGGAAATG	GTCCAACAGA	3480
AAAAAGAGGG	CCCCATTCCA	CCAGATGCAC	AAAAATCCAGA	TATGTCGTTT	TTTAAGATGC	3540
TATTCTTGCC	AGAATCAGCA	AGGTGGATAC	AAAGGACTCA	TGGAAAGAAC	TCTCTGAACT	3600
CTGGGCAAGG	CCCCAGTCCA	AAGCAATTAG	TATCCTTAGG	ACCAGAAAAA	TCTGTGGAAG	3660
GTCAGAAATTT	CTTGTCTGAG	AAAAACAAAG	TGGTAGTAGG	AAAGGGTGAA	TTTACAAAGG	3720
ACGTAGGACT	CAAAAGAGATG	GTTTTTCCAA	GCAGCAGAAA	CCTATTTCTT	ACTAACTTGG	3780
ATAATTTACA	TGAAAAATAAT	ACACACAATC	AAGAAAAAAA	AATTCAGGAA	GAAATAGAAA	3840

AGAAGGAAAC	ATTAATCCAA	GAGAATGTAG	TTTTGCCTCA	GATACATACA	GTGACTGGCA	3900
CTAAGAATTT	CATGAAGAAC	CTTTTCTTAC	TGAGCACTAG	GCAAAATGTA	GAAGGTTTAT	3960
ATGACGGGGC	ATATGCTCCA	GTACTTCAAG	ATTTTAGGTC	ATTAAATGAT	TCAACAAATA	4020
GAACAAAGAA	ACACACAGCT	CATTTCTCAA	AAAAAGGGGA	GGAAGAAAAC	TTGGAAGGCT	4080
TGGGAAATCA	AACCAAGCAA	ATTGTAGAGA	AATATGCATG	CACCACAAGG	ATATCTCCTA	4140
ATACAAGCCA	GCAGAAATTT	GTCACGCAAC	GTAGTAAGAG	AGCTTTGAAA	CAATTCAGAC	4200
TCCCAC TAGA	AGAAACAGAA	CTTGAAAAAA	GGATAATTGT	GGATGACACC	TCAACCCAGT	4260
GGTCCAAAAA	CATGAAACAT	TTGACCCCGA	GCACCCTCAC	ACAGATAGAC	TACAATGAGA	4320
AGGAGAAAGG	GGCCATTACT	CAGTCTCCCT	TATCAGATTG	CCTTACGAGG	AGTCATAGCA	4380
TCCCTCAAGC	AAATAGATCT	CCATTACCCA	TTGCAAAGGT	ATCATCATTT	CCATCTATTA	4440
GACCTATATA	TCTGACCAGG	GTCCATTATCC	AAGACAACCT	TTCTCATCTT	CCAGCAGCAT	4500
CTTATAGAAA	GAAAGATTCT	GGGGTCCAAG	AAAGCAGTCA	TTTCTTACAA	GGAGCCAAAA	4560
AAAATAACCT	TTCTTTAGCC	ATTCTAACCT	TGGAGATGAC	TGGTGATCAA	AGAGAGGTTG	4620
GCTCCCTGGG	GACAAGTGCC	ACAAATTCAG	TCACATACAA	GAAAGTTGAG	AACACTGTTC	4680
TCCCGAAACC	AGACTTGCCC	AAAACATCTG	GCAAAGTTGA	ATTGCTTCCA	AAAGTTCACA	4740
TTTATCAGAA	GGACCTATTC	CCTACGGAAA	CTAGCAATGG	GTCTCCTGGC	CATCTGGATC	4800
TCGTGGAAGG	GAGCCTTCTT	CAGGGAACAG	AGGGAGCGAT	TAAGTGGAAT	GAAGCAACA	4860
GACCTGGAAA	AGTTCCCTTT	CTGAGAGTAG	CAACAGAAAG	CTCTGCAAAG	ACTCCCTCCA	4920
AGCTATTGGA	TCCCTTTGCT	TGGGATAACC	ACTATGTTAC	TCAGATACCA	AAAGAAGAGT	4980
GGAAATCCCA	AGAGAAGTCA	CCAGAAAAAA	CAGCTTTTAA	GAAAAAGGAT	ACCATTTTGT	5040
CCCTGAACGC	TTGTGAAAGC	AATCATGCAA	TAGCAACAA	AAATGAGGGA	CAAAATAAGC	5100
CCGAAATAGA	AGTCACCTGG	GCAAAGCAAG	GTAGGACTGA	AAGGCTGTGC	TCTCAAAACC	5160
CACCAGTCTT	GAAACGCCAT	CAACGGGAAA	TAACTCGTAC	TACTCTTCAG	TCAGATCAAG	5220
AGGAAATGTA	CTATGATGAT	ACCATATCAG	TTGAAATGAA	GAAGGAAGAT	TTTGACATTT	5280
ATGATGAGGA	TGAAAATCAG	AGCCCCCGCA	GCTTTCAAAA	GAAAACACGA	CACATATTTA	5340
TTGCTGCAGT	GGAGAGGCTC	TGGGATTATG	GGATGAGTAG	CTCCCCACAT	GTTCTAAGAA	5400
ACAGGGCTCA	GAGTGGCAGT	GTCCCTCAGT	TCAAAGAAAGT	TGTTTTCCAG	GAATTTACTG	5460
ATGGCTCCTT	TACTCAGCCC	TTATACCGTG	GAGAACTAAA	TGAACATTTG	GGACTCCTGG	5520
GGCCATATAT	AAGAGCAGAA	GTTGAAGATA	ATATCATGGT	AACTTTCAGA	AATCAGGCCT	5580
CTCGTCCCTA	TTCTTTCTAT	TCTAGCCTTA	TTTCTTATGA	GGAAGATCAG	AGGCAAGGAG	5640
CAGAACCTAG	AAAAAACTTT	GTCAAGCCTA	ATGAAACCAA	AACTTACTTT	TGGAAAGTGC	5700
AACATCATAT	GGCACCCACT	AAAGATGAGT	TTGACTGCAA	AGCCTGGGCT	TATTTCTCTG	5760
ATGTTGACCT	GGAAAAAGAT	GTGCACTCAG	GCTTATTTGG	ACCCCTTCTG	GTCTGCCACA	5820
CTAACACACT	GAACCCCTGCT	CATGGGAGAC	AAGTGACAGT	ACAGGAATTT	GCTCTGTTTT	5880
TCACCATCTT	TGATGAGACC	AAAAGCTGGT	ACTTCACTGA	AAATATGGAA	AGAAACTGCA	5940
GGGCTCCCTG	CAATATCCAG	ATGGAAGATC	CCACTTTTAA	AGAGAATTAT	CGCTTCCATG	6000
CAATCAATGG	CTACATAATG	GATACACTAC	CTGGCTTAGT	AATGGCTCAG	GATCAAAGGA	6060
TTTCGATGGTA	TCTGCTCAGC	ATGGGCAGCA	ATGAAAACAT	CCATTCTATT	CATTTTCAGTG	6120
GACATGTGTT	CACTGTACGA	AAAAAAGAGG	AGTATAAAAT	GGCACTGTAC	AATCTCTATC	6180
CAGGTGTTTT	TGAGACAGTG	GAAATGTTAC	CATCCAAAGC	TGGAATTTGG	CGGGTGGAA	6240
GCCTTATTGG	CGAGCATCTA	CATGCTGGGA	TGAGCACACT	TTTTCTGGTG	TACAGCAATA	6300
AGTGTACAGC	TCCCCTGGGA	ATGGCTTCTG	GACACATTAG	AGATTTTCAG	ATTACAGCTT	6360
CAGGACAATA	TGGACAGTGG	GCCCCAAAGC	TGGCCAGACT	TCATTATTCC	GGATCAATCA	6420
ATGCCTGGAG	CACCAAGGAG	CCCTTTTCTT	GGATCAAGGT	GGATCTGTTG	GCACCAATGA	6480
TTATTCACGG	CATCAAGACC	CAGGGTGCCC	GTCAGAAAGT	CTCCAGCCTC	TACATCTCTC	6540
AGTTTATCAT	CATGTATAGT	CTTGATGGGA	AGAAGTGGCA	GACTTATCGA	GGAAATTTCA	6600
CTGGAACCTT	AATGGTCTTC	TTTGGCAATG	TGGATTCATC	TGGGATAAAA	CACAATATTT	6660
TTAACCCCTC	AATTATTGCT	CGATACATCC	GTTTGCACCC	AACTCATTAT	AGCATTCGCA	6720
GCACTCTTCG	CATGGAGTTG	ATGGGCTGTG	ATTTAAATAG	TTGCAGCATG	CCATTGGGAA	6780
TGGAGAGTAA	AGCAATATCA	GATGCACAGA	TTACTGCTTC	ATCCTACTTT	ACCAATATGT	6840
TTGCCACCTG	GTCTCCTTCA	AAAGCTCGAC	TTCACTCCA	AGGGAGGAGT	AATGCCTGGA	6900
GACCTCAGGT	GAATAATCCA	AAAGAGTGGC	TGCAAGTGGA	CTTCCAGAAG	ACAATGAAAG	6960
TCACAGGAGT	AACTACTCAG	GGAGTAAAA	CTCTGCTTAC	CAGCATGTAT	GTGAAGGAGT	7020
TCCTCATCTC	CAGCAGTCAA	GATGGCCATC	AGTGGACTCT	CTTTTTCAG	AATGGCAAAG	7080
TAAAGCTTTT	TGAGGAAAA	CAAGACTCCT	TCACACTGT	GGTGAACCTC	CTAGACCCAC	7140
CGTTACTGAC	TCGCTACCTT	CGAATTCACC	CCCAGAGTTG	GGTGCACCAG	ATTGCCCTGA	7200
GGATGGAGGT	TCTGGGCTGC	GAGGCACAGG	ACCTCTACTG	AGGGTGGCCA	CTGCAGCACC	7260
TGCCACTGCC	GTCACCTCTC	CCTCCTCAGC	TCCAGGCGAG	TGTCCCTCCC	TGGCTTGCCCT	7320
TCTACCTTTG	TGCTAAATCC	TAGCAGACAC	TGCCCTGAAG	CCTCCTGAAT	TAACTATCAT	7380
CAGTCTTGCA	TTTCTTTGGT	GGGGGGCCAG	GAGGGTGCAT	CCAATTTAAC	TTAACTCTTA	7440
CCTATTTTCT	GCAGCTGCTC	CCAGATTACT	CCTTCTTCC	AATATAACTA	GGCAAAAAGA	7500

AGTGAGGAGA	AACCTGCATG	AAAGCATTTCT	TCCCTGAAAA	GTTAGGCCTC	TCAGAGTCAC	7560
CACTTCCTCT	GTGTAGAAAA	AACTATGTGA	TGAAACTTTG	AAAAAGATAT	TTATGATGTT	7620
AACATTTTTCAG	GTTAAGCCTC	ATACGTTTAA	AATAAACTC	TCAGTTGTTT	ATTATCCTGA	7680
TCAAGCATGG	AACAAAGCAT	GTTTCAGGAT	CAGATCAATA	CAATCTTGGA	GTCAAAAGGC	7740
AAATCATTTG	GACAATCTGC	AAAATGGAGA	GAATACAATA	ACTACTACAG	TAAAGTCTGT	7800
TTCTGCTTCC	TTACACATAG	ATATAATTAT	GTTATTTAGT	CATTATGAGG	GGCACATTCT	7860
TATCTCCAAA	ACTAGCATTC	TTAAACTGAG	AATTATAGAT	GGGGTTCAAG	AATCCCTAAG	7920
TCCCCTGAAA	TTATATAAGG	CATTCTGTAT	AAATGCAAAAT	GTGCATTTTT	CTGACGAGTG	7980
TCCATAGATA	TAAAGCCATT	TGGTCTTAAT	TCTGACCAAT	AAAAAAATAA	GTCAGGAGGA	8040
TGCAATTGTT	GAAAGCTTTG	AAATAAAATA	ACAATGTCTT	CTTGAAATTT	GTGATGGCCA	8100
AGAAAGAAAA	TGATGATGAC	ATTAGGCTTC	TAAAGGACAT	ACATTTAATA	TTTCTGTGGA	8160
AATATGAGGA	AAATCCATGG	TTATCTGAGA	TAGGAGATAC	AAACTTTGTA	ATTCTAATAA	8220
TGCACTCAGT	TTACTCTCTC	CCTCTACTAA	TTTCCTGCTG	AAAATAACAC	AACAAAAATG	8280
TAACAGGGGA	AATTATATAC	CGTGACTGAA	AACTAGAGTC	CTACTTACAT	AGTTGAAATA	8340
TCAAGGAGGT	CAGAAGAAAA	TTGGACTGGT	GAAAACAGAA	AAAACACTCC	AGTCTGCCAT	8400
ATCACCACAC	AATAGGATCC	CCCTTCTTGC	CCTCCACCCC	CATAAGATTG	TGAAGGGTTT	8460
ACTGCTCCTT	CCATCTGCCT	GACCCCTTCA	CTATGACTAC	ACAGAATCTC	CTGATAGTAA	8520
AGGGGGCTGG	AGGCAAGGAT	AAGTTATAGA	GCAGTTGGAG	GAAGCATCCA	AAGATTGCAA	8580
CCCAGGGCAA	ATGGAAAACA	GGAGATCCTA	ATAGGAAAGA	AAAATGGATC	CCAATCTGAG	8640
AAAAGGCAAA	AGAATGGCTA	CTTTTTTCTA	TGCTGGAGTA	TTTTCTAATA	ATCCTGCTTG	8700
ACCCTTATCT	GACCTCTTTG	GAAACTATAA	CATAGCTGTC	ACAGTATAGT	CACAATCCAC	8760
AAATGATGCA	GGTGCAAATG	GTTTATAGCC	CTGTGAAGTT	CTTAAAGTTT	AGAGGCTAAC	8820
TTACAGAAAT	GAATAAGTTG	TTTTGTTTTA	TAGCCCGGTA	GAGGAGTTAA	CCCCAAAGGT	8880
GATATGGTTT	TATTTCTCTG	TATGTTTAAAC	TTAATAATCT	TATTTTGGCA	TTCTTTTCCC	8940
ATTGACTATA	TACATCTCTA	TTTCTCAAAAT	GTTTCATGGAA	CTAGCTCTTT	TATTTTCTCTG	9000
CTGGTTTCTT	CAGTAATGAG	TTAAATAAAA	CATTGACACA	TACAAAAAAA	AAAAAAAATA	9060
AAAAAAAATA	AAAAAAAATA					9080

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2351 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

Met	Gln	Ile	Glu	Leu	Ser	Thr	Cys	Phe	Phe	Leu	Cys	Leu	Leu	Arg	Phe	1	5	10	15
Cys	Phe	Ser	Ala	Thr	Arg	Arg	Tyr	Tyr	Leu	Gly	Ala	Val	Glu	Leu	Ser	20	25	30	
Trp	Asp	Tyr	Met	Gln	Ser	Asp	Leu	Gly	Glu	Leu	Pro	Val	Asp	Ala	Arg	35	40	45	
Phe	Pro	Pro	Arg	Val	Pro	Lys	Ser	Phe	Pro	Phe	Asn	Thr	Ser	Val	Val	50	55	60	
Tyr	Lys	Lys	Thr	Leu	Phe	Val	Glu	Phe	Thr	Asp	His	Leu	Phe	Asn	Ile	65	70	75	80
Ala	Lys	Pro	Arg	Pro	Pro	Trp	Met	Gly	Leu	Leu	Gly	Pro	Thr	Ile	Gln	85	90	95	
Ala	Glu	Val	Tyr	Asp	Thr	Val	Val	Ile	Thr	Leu	Lys	Asn	Met	Ala	Ser	100	105	110	
His	Pro	Val	Ser	Leu	His	Ala	Val	Gly	Val	Ser	Tyr	Trp	Lys	Ala	Ser	115	120	125	

Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp
 130 135 140
 Asp Lys Val Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu
 145 150 155 160
 Lys Glu Asn Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser
 165 170 175
 Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile
 180 185 190
 Gly Ala Leu Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr
 195 200 205
 Gln Thr Leu His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly
 210 215 220
 Lys Ser Trp His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp
 225 230 235 240
 Ala Ala Ser Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr
 245 250 255
 Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val
 260 265 270
 Tyr Trp His Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile
 275 280 285
 Phe Leu Glu Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser
 290 295 300
 Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met
 305 310 315 320
 Asp Leu Gly Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His
 325 330 335
 Asp Gly Met Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro
 340 345 350
 Gln Leu Arg Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp
 355 360 365
 Leu Thr Asp Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser
 370 375 380
 Pro Ser Phe Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr
 385 390 395 400
 Trp Val His Tyr Ile Ala Ala Glu Glu Glu Asp Trp Asp Tyr Ala Pro
 405 410 415
 Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn
 420 425 430
 Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met
 435 440 445
 Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu

450	455	460
Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu		
465	470	475 480
Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro		
	485	490 495
His Gly Ile Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys		
	500	505 510
Gly Val Lys His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe		
	515	520 525
Lys Tyr Lys Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp		
	530	535 540
Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg		
	545	550 555 560
Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu		
	565	570 575
Ser Val Asp Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val		
	580	585 590
Ile Leu Phe Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu		
	595	600 605
Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp		
	610	615 620
Pro Glu Phe Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val		
	625	630 635 640
Phe Asp Ser Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp		
	645	650 655
Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe		
	660	665 670
Ser Gly Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr		
	675	680 685
Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro		
	690	695 700
Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly		
	705	710 715 720
Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp		
	725	730 735
Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys		
	740	745 750
Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser Gln Asn Ser Arg His Pro		
	755	760 765
Ser Thr Arg Gln Lys Gln Phe Asn Ala Thr Thr Ile Pro Glu Asn Asp		
	770	775 780

Ile Glu Lys Thr Asp Pro Trp Phe Ala His Arg Thr Pro Met Pro Lys
 785 790 795 800
 Ile Gln Asn Val Ser Ser Ser Asp Leu Leu Met Leu Leu Arg Gln Ser
 805 810 815
 Pro Thr Pro His Gly Leu Ser Leu Ser Asp Leu Gln Glu Ala Lys Tyr
 820 825 830
 Glu Thr Phe Ser Asp Asp Pro Ser Pro Gly Ala Ile Asp Ser Asn Asn
 835 840 845
 Ser Leu Ser Glu Met Thr His Phe Arg Pro Gln Leu His His Ser Gly
 850 855 860
 Asp Met Val Phe Thr Pro Glu Ser Gly Leu Gln Leu Arg Leu Asn Glu
 865 870 875 880
 Lys Leu Gly Thr Thr Ala Ala Thr Glu Leu Lys Lys Leu Asp Phe Lys
 885 890 895
 Val Ser Ser Thr Ser Asn Asn Leu Ile Ser Thr Ile Pro Ser Asp Asn
 900 905 910
 Leu Ala Ala Gly Thr Asp Asn Thr Ser Ser Leu Gly Pro Pro Ser Met
 915 920 925
 Pro Val His Tyr Asp Ser Gln Leu Asp Thr Thr Leu Phe Gly Lys Lys
 930 935 940
 Ser Ser Pro Leu Thr Glu Ser Gly Gly Pro Leu Ser Leu Ser Glu Glu
 945 950 955 960
 Asn Asn Asp Ser Lys Leu Leu Glu Ser Gly Leu Met Asn Ser Gln Glu
 965 970 975
 Ser Ser Trp Gly Lys Asn Val Ser Ser Thr Glu Ser Gly Arg Leu Phe
 980 985 990
 Lys Gly Lys Arg Ala His Gly Pro Ala Leu Leu Thr Lys Asp Asn Ala
 995 1000 1005
 Leu Phe Lys Val Ser Ile Ser Leu Leu Lys Thr Asn Lys Thr Ser Asn
 1010 1015 1020
 Asn Ser Ala Thr Asn Arg Lys Thr His Ile Asp Gly Pro Ser Leu Leu
 1025 1030 1035 1040
 Ile Glu Asn Ser Pro Ser Val Trp Gln Asn Ile Leu Glu Ser Asp Thr
 1045 1050 1055
 Glu Phe Lys Lys Val Thr Pro Leu Ile His Asp Arg Met Leu Met Asp
 1060 1065 1070
 Lys Asn Ala Thr Ala Leu Arg Leu Asn His Met Ser Asn Lys Thr Thr
 1075 1080 1085
 Ser Ser Lys Asn Met Glu Met Val Gln Gln Lys Lys Glu Gly Pro Ile
 1090 1095 1100

Pro Pro Asp Ala Gln Asn Pro Asp Met Ser Phe Phe Lys Met Leu Phe
 1105 1110 1115 1120
 Leu Pro Glu Ser Ala Arg Trp Ile Gln Arg Thr His Gly Lys Asn Ser
 1125 1130 1135
 Leu Asn Ser Gly Gln Gly Pro Ser Pro Lys Gln Leu Val Ser Leu Gly
 1140 1145 1150
 Pro Glu Lys Ser Val Glu Gly Gln Asn Phe Leu Ser Glu Lys Asn Lys
 1155 1160 1165
 Val Val Val Gly Lys Gly Glu Phe Thr Lys Asp Val Gly Leu Lys Glu
 1170 1175 1180
 Met Val Phe Pro Ser Ser Arg Asn Leu Phe Leu Thr Asn Leu Asp Asn
 1185 1190 1195 1200
 Leu His Glu Asn Asn Thr His Asn Gln Glu Lys Lys Ile Gln Glu Glu
 1205 1210 1215
 Ile Glu Lys Lys Glu Thr Leu Ile Gln Glu Asn Val Val Leu Pro Gln
 1220 1225 1230
 Ile His Thr Val Thr Gly Thr Lys Asn Phe Met Lys Asn Leu Phe Leu
 1235 1240 1245
 Leu Ser Thr Arg Gln Asn Val Glu Gly Ser Tyr Asp Gly Ala Tyr Ala
 1250 1255 1260
 Pro Val Leu Gln Asp Phe Arg Ser Leu Asn Asp Ser Thr Asn Arg Thr
 1265 1270 1275 1280
 Lys Lys His Thr Ala His Phe Ser Lys Lys Gly Glu Glu Glu Asn Leu
 1285 1290 1295
 Glu Gly Leu Gly Asn Gln Thr Lys Gln Ile Val Glu Lys Tyr Ala Cys
 1300 1305 1310
 Thr Thr Arg Ile Ser Pro Asn Thr Ser Gln Gln Asn Phe Val Thr Gln
 1315 1320 1325
 Arg Ser Lys Arg Ala Leu Lys Gln Phe Arg Leu Pro Leu Glu Glu Thr
 1330 1335 1340
 Glu Leu Glu Lys Arg Ile Ile Val Asp Asp Thr Ser Thr Gln Trp Ser
 1345 1350 1355 1360
 Lys Asn Met Lys His Leu Thr Pro Ser Thr Leu Thr Gln Ile Asp Tyr
 1365 1370 1375
 Asn Glu Lys Glu Lys Gly Ala Ile Thr Gln Ser Pro Leu Ser Asp Cys
 1380 1385 1390
 Leu Thr Arg Ser His Ser Ile Pro Gln Ala Asn Arg Ser Pro Leu Pro
 1395 1400 1405
 Ile Ala Lys Val Ser Ser Phe Pro Ser Ile Arg Pro Ile Tyr Leu Thr
 1410 1415 1420
 Arg Val Leu Phe Gln Asp Asn Ser Ser His Leu Pro Ala Ala Ser Tyr

1425	1430	1435	1440
Arg Lys Lys Asp Ser Gly Val Gln Glu Ser Ser His Phe Leu Gln Gly	1445	1450	1455
Ala Lys Lys Asn Asn Leu Ser Leu Ala Ile Leu Thr Leu Glu Met Thr	1460	1465	1470
Gly Asp Gln Arg Glu Val Gly Ser Leu Gly Thr Ser Ala Thr Asn Ser	1475	1480	1485
Val Thr Tyr Lys Lys Val Glu Asn Thr Val Leu Pro Lys Pro Asp Leu	1490	1495	1500
Pro Lys Thr Ser Gly Lys Val Glu Leu Leu Pro Lys Val His Ile Tyr	1505	1510	1515
Gln Lys Asp Leu Phe Pro Thr Glu Thr Ser Asn Gly Ser Pro Gly His	1525	1530	1535
Leu Asp Leu Val Glu Gly Ser Leu Leu Gln Gly Thr Glu Gly Ala Ile	1540	1545	1550
Lys Trp Asn Glu Ala Asn Arg Pro Gly Lys Val Pro Phe Leu Arg Val	1555	1560	1565
Ala Thr Glu Ser Ser Ala Lys Thr Pro Ser Lys Leu Leu Asp Pro Leu	1570	1575	1580
Ala Trp Asp Asn His Tyr Gly Thr Gln Ile Pro Lys Glu Glu Trp Lys	1585	1590	1595
Ser Gln Glu Lys Ser Pro Glu Lys Thr Ala Phe Lys Lys Lys Asp Thr	1605	1610	1615
Ile Leu Ser Leu Asn Ala Cys Glu Ser Asn His Ala Ile Ala Ala Ile	1620	1625	1630
Asn Glu Gly Gln Asn Lys Pro Glu Ile Glu Val Thr Trp Ala Lys Gln	1635	1640	1645
Gly Arg Thr Glu Arg Leu Cys Ser Gln Asn Pro Pro Val Leu Lys Arg	1650	1655	1660
His Gln Arg Glu Ile Thr Arg Thr Thr Leu Gln Ser Asp Gln Glu Glu	1665	1670	1675
Ile Asp Tyr Asp Asp Thr Ile Ser Val Glu Met Lys Lys Glu Asp Phe	1685	1690	1695
Asp Ile Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe Gln Lys	1700	1705	1710
Lys Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp Asp Tyr	1715	1720	1725
Gly Met Ser Ser Ser Pro His Val Leu Arg Asn Arg Ala Gln Ser Gly	1730	1735	1740
Ser Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr Asp Gly	1745	1750	1755
			1760

Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu His Leu Gly
 1765 1770 1775
 Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp Asn Ile Met Val
 1780 1785 1790
 Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser Phe Tyr Ser Ser Leu
 1795 1800 1805
 Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu Pro Arg Lys Asn
 1810 1815 1820
 Phe Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys Val Gln His
 1825 1830 1835 1840
 His Met Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys Ala Trp Ala Tyr
 1845 1850 1855
 Phe Ser Asp Val Asp Leu Glu Lys Asp Val His Ser Gly Leu Ile Gly
 1860 1865 1870
 Pro Leu Leu Val Cys His Thr Asn Thr Leu Asn Pro Ala His Gly Arg
 1875 1880 1885
 Gln Val Thr Val Gln Glu Phe Ala Leu Phe Phe Thr Ile Phe Asp Glu
 1890 1895 1900
 Thr Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg Asn Cys Arg Ala
 1905 1910 1915 1920
 Pro Cys Asn Ile Gln Met Glu Asp Pro Thr Phe Lys Glu Asn Tyr Arg
 1925 1930 1935
 Phe His Ala Ile Asn Gly Tyr Ile Met Asp Thr Leu Pro Gly Leu Val
 1940 1945 1950
 Met Ala Gln Asp Gln Arg Ile Arg Trp Tyr Leu Leu Ser Met Gly Ser
 1955 1960 1965
 Asn Glu Asn Ile His Ser Ile His Phe Ser Gly His Val Phe Thr Val
 1970 1975 1980
 Arg Lys Lys Glu Glu Tyr Lys Met Ala Leu Tyr Asn Leu Tyr Pro Gly
 1985 1990 1995 2000
 Val Phe Glu Thr Val Glu Met Leu Pro Ser Lys Ala Gly Ile Trp Arg
 2005 2010 2015
 Val Glu Cys Leu Ile Gly Glu His Leu His Ala Gly Met Ser Thr Leu
 2020 2025 2030
 Phe Leu Val Tyr Ser Asn Lys Cys Gln Thr Pro Leu Gly Met Ala Ser
 2035 2040 2045
 Gly His Ile Arg Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr Gly Gln
 2050 2055 2060
 Trp Ala Pro Lys Leu Ala Arg Leu His Tyr Ser Gly Ser Ile Asn Ala
 2065 2070 2075 2080

Trp Ser Thr Lys Glu Pro Phe Ser Trp Ile Lys Val Asp Leu Leu Ala
 2085 2090 2095
 Pro Met Ile Ile His Gly Ile Lys Thr Gln Gly Ala Arg Gln Lys Phe
 2100 2105 2110
 Ser Ser Leu Tyr Ile Ser Gln Phe Ile Ile Met Tyr Ser Leu Asp Gly
 2115 2120 2125
 Lys Lys Trp Gln Thr Tyr Arg Gly Asn Ser Thr Gly Thr Leu Met Val
 2130 2135 2140
 Phe Phe Gly Asn Val Asp Ser Ser Gly Ile Lys His Asn Ile Phe Asn
 2145 2150 2155 2160
 Pro Pro Ile Ile Ala Arg Tyr Ile Arg Leu His Pro Thr His Tyr Ser
 2165 2170 2175
 Ile Arg Ser Thr Leu Arg Met Glu Leu Met Gly Cys Asp Leu Asn Ser
 2180 2185 2190
 Cys Ser Met Pro Leu Gly Met Glu Ser Lys Ala Ile Ser Asp Ala Gln
 2195 2200 2205
 Ile Thr Ala Ser Ser Tyr Phe Thr Asn Met Phe Ala Thr Trp Ser Pro
 2210 2215 2220
 Ser Lys Ala Arg Leu His Leu Gln Gly Arg Ser Asn Ala Trp Arg Pro
 2225 2230 2235 2240
 Gln Val Asn Asn Pro Lys Glu Trp Leu Gln Val Asp Phe Gln Lys Thr
 2245 2250 2255
 Met Lys Val Thr Gly Val Thr Thr Gln Gly Val Lys Ser Leu Leu Thr
 2260 2265 2270
 Ser Met Tyr Val Lys Glu Phe Leu Ile Ser Ser Ser Gln Asp Gly His
 2275 2280 2285
 Gln Trp Thr Leu Phe Phe Gln Asn Gly Lys Val Lys Val Phe Gln Gly
 2290 2295 2300
 Asn Gln Asp Ser Phe Thr Pro Val Val Asn Ser Leu Asp Pro Pro Leu
 2305 2310 2315 2320
 Leu Thr Arg Tyr Leu Arg Ile His Pro Gln Ser Trp Val His Gln Ile
 2325 2330 2335
 Ala Leu Arg Met Glu Val Leu Gly Cys Glu Ala Gln Asp Leu Tyr
 2340 2345 2350

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4832 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

CTCGAGCTAA AGATATTTTA GAGAAGAATT AACCTTTTGC TTCTCCAGTT GAACATTTGT

AGCAATAAGT	CATGCAAATA	GAGCTCTCCA	CCTGCTTCTT	TCTGTGCCTT	TTGCGATTCT	120
GCTTTAGTGC	CACCAGAAGA	TACTACCTGG	GTGCAGTGG	ACTGTCTATG	GACTATATGC	180
AAAGTGATCT	CGGTGAGCTG	CCTGTGGACG	CAAGATTTC	TCCTAGAGTG	CCAAAATCTT	240
TTCCATTCAA	CACCTCAGTC	GTGTACAAAA	AGACTCTGTT	TGTAGAATTC	ACGGATCACC	300
TTTTCAACAT	CGCTAAGCCA	AGGCCACCTT	GGATGGGTCT	GCTAGGTCCT	ACCATCCAGG	360
CTGAGGTTTA	TGATACAGTG	GTCATTACAC	TTAAGAACAT	GGCTTCCCAT	CCTGTCTAGT	420
TCATGCTGT	TGGTGTATCC	TACTGGAAAG	CTTCTGAGGG	AGCTGAATAT	GATGATCAGA	480
CCAGTCAAAAG	GGAGAAAGAA	GATGATAAAG	TCTTCCCTGG	TGGAAGCCAT	ACATATGTCT	540
GGCAGGTCTT	GAAAGAGAAT	GGTCCAATGG	CCTCTGACCC	ACTGTGCCTT	ACCTACTCAT	600
ATCTTTCTCA	TGTGGACCTG	GTAAAAGACT	TGAATTCAGG	CCTCATTGGA	GCCCTACTAG	660
TATGTAGAGA	AGGGAGTCTG	GCCAAGGAAA	AGACACAGAC	CTTGACACAA	TTTATACTAC	720
TTTTTGCTGT	ATTTTGATGAA	GGGAAAAGTT	GGCACTCAGA	AACAAAGAAC	TCCTTGATGC	780
AGGATAGGGA	TGCTGCATCT	GCTCGGGCCT	GGCCTAAAAT	GCACACAGTC	AATGGTTATG	840
TAAACAGGTC	TCTGCCAGGT	CTGATTGGAT	GCCACAGGAA	ATCAGTCTAT	TGGCATGTGA	900
TTGGAATGGG	CACCACTCCT	GAAGTGCAC	CAATATTTCCT	CGAAGGTCAC	ACATTTCTTG	960
TGAGGAACCA	TCGCCAGGCG	TCCTTGGA	TCTCGCCAAT	AACTTTCTCT	ACTGCTCAAA	1020
CACTCTTGAT	GGACCTTGGA	CAGTTTCTAC	TGTTTGTGCA	TATCTCTTCC	CACCAACATG	1080
ATGGCATGGA	AGCTTATGTC	AAAGTAGACA	GCTGTCCAGA	GGAACCCCAA	CTACGAATGA	1140
AAAATAATGA	AGAAGCGGAA	GACTATGATG	ATGATCTTAC	TGATTCTGAA	ATGGATGTGG	1200
TCAGGTTTGA	TGATGACAAC	TCTCCTTCCT	TTATCCAAAT	TCGCTCAGTT	GCCAAGAAGC	1260
ATCCTAAAC	TTGGGTACAT	TACATTGCTG	CTGAAGAGGA	GGACTGGGAC	TATGCTCCCT	1320
TAGTCTTCGC	CCCCGATGAC	AGAAGTTATA	AAAGTCAATA	TTTGAACAAT	GGCCCTCAGC	1380
GGATTGGTAG	GAAGTACAAA	AAAGTCCGAT	TTATGGCATA	CACAGATGAA	ACCTTTAAGA	1440
CTCGTGAAGC	TATTCAGCAT	GAATCAGGAA	TCTTGGGACC	TTTACTTTAT	GGGGAAGTTG	1500
GAGACACACT	GTTGATTATA	TTAAGAATC	AAGCAAGCAG	ACCATATAAC	ATCTACCCTC	1560
ACGGAATCAC	TGATGTCCGT	CCTTTGTATT	CAAGGAGATT	ACCAAAAGGT	GTAAAACATT	1620
TGAAGGATTT	TCCAATTCTG	CCAGGAGAAA	TATTCAAATA	TAAATGGACA	GTGACTGTAG	1680
AAGATGGGCC	AACATAATCA	GATCCTCGGT	GCCTGACCCG	CTATTACTCT	AGTTTCGTTA	1740
ATATGGAGAG	AGATCTAGCT	TCAGGACTCA	TTGGCCCTCT	CCTCATCTGC	TACAAAGAAT	1800
CTGTAGATCA	AAGAGGAAAC	CAGATAATGT	CAGACAAGAG	GAATGTCATC	CTGTTTTCTG	1860
TATTTGATGA	GAACCGAAGC	TGGTACCTCA	CAGAGAATAT	ACAACGCTTT	CTCCCCAATC	1920
CAGCTGGAGT	GCAGCTTGAG	GATCCAGAGT	TCCAAGCCTC	CAACATCATG	CACAGCATCA	1980
ATGGCTATGT	TTTTGATAGT	TTGAGTGTGT	QAGTTTGT	GCATGAGGTG	GCATAGGTTG	2040
ACATTCTAAG	CATTGGAGCA	CAGACTGACT	TCCTTTCTGT	CTTCTTCTCT	GGATATACCT	2100
TCAAACACAA	AATGGTCTAT	GAAGACACAC	TCACCTATT	CCCATTCTCA	GGAGAACTG	2160
TCTTCATGTC	GATGGAAAAC	CCAGGTCTAT	GGATTCTGGG	GTGCCACAAC	TCAGACTTTC	2220
GGAACAGAGG	CATGACCGCC	TTACTGAAGG	TTTCTAGTTG	TGACAAGAAC	ACTGGTGATT	2280
ATTACGAGGA	CAGTTATGAA	GATATTTTCA	CATACTTGCT	GAGTAAAAAC	AATGCCATTG	2340
AACCAAGAAG	CTTCTCCAG	AACCCACCA	TCTTGAAACG	CCATCAACGG	GAAATAACTC	2400
GTACTACTCT	TCAGTCAGAT	CAAGAGGAAA	TTGACTATGA	TGATACCATA	TCAGTTGAAA	2460
TGAAGAAGGA	AGATTTTGAC	ATTTATGATG	AGGATGAAAA	TCAGAGCCCC	CGCAGCTTTC	2520
AAAAGAAAAC	ACGACACTAT	TTTATGCTG	CAGTGGAGAG	GCTCTGGGAT	TATGGGATGA	2580
GTAGCTCCCC	ACATGTTCTA	AGAAACAGGG	CTCAGAGTGG	CAGTGTCCCT	CAGTTCAAGA	2640
AAGTTGTTTT	CCAGGAATTT	ACTGATGGCT	CCTTTACTCA	GCCCTTATAC	CGTGGAGAAC	2700
TAAATGAACA	TTTGGGACTC	CTGGGGCCAT	ATATAAGAGC	AGAAGTTGAA	GATAATATCA	2760
TGGTAACTTT	CAGAAATCAG	GCCTCTCGTC	CCTATTCTCT	CTATTCTAGC	CTTATTTCTT	2820
ATGAGGAAGA	TCAGAGGCAA	GGAGCAGAAC	CTAGAAAAAA	CCTTGTCAGG	CCTAATGAAA	2880
CCAAAACCTA	CTTTTGAAAA	GTGCAACATC	ATATGGCACC	CACTAAAAGAT	GAGTTTGACT	2940
GCAAAGCCTG	GGCTTATTTT	TCTGATGTTG	ACCTGGAAAA	AGATGTGCAC	TCAGGCCTGA	3000
TTGGACCCCT	TCTGGTCTGC	CACACTAACA	CACTGAACCC	TGCTCATGGG	AGACAAGTGA	3060
CAGTACAGGA	ATTTGCTCTG	TTTTTCAACA	TCTTTGATGA	GACCAAAAGC	TGGTACTTCA	3120
CTGAAAAAT	GGAAAGAAAC	TGCAGGGCTC	CCTGCAATAT	CCAGATGGAA	GATCCCCTT	3180
TTAAAGAGAA	TTATCGCTTC	CATGCAATCA	ATGGCTACAT	AATGGATACA	CTACCTGGCT	3240
TAGTAATGGC	TCAGGATCAA	AGGATTGATG	GGTATCTGCT	CAGCATGGGC	AGCAATGAAA	3300
ACATCCATTC	TATTCATTTT	AGTGGACATG	TGTTCACTGT	ACGAAAAAAA	GAGGAGTATA	3360
AAATGGCACT	GTACAAATCT	TATCCAGGTG	TTTTTGAGAC	AGTGGAAATG	TTACCATCCA	3420
AAGCTGGAAT	TTGGCGGGTG	GAATGCCCTT	TTGGCGAGCA	TCTACATGCT	GGGATGAGCA	3480
CACTTTTTTCT	GGTGTACAGC	AATAAGTGTC	AGACTCCCCCT	GGGAATGGCT	TCTGGACACA	3540
TTAGAGATTT	TCAGATTACA	GCTTCAGGAC	AATATGGACA	GTGGGCCCCA	AAGCTGGCCA	3600
GACTTCATTA	TTCCGGATCA	ATCAATGCCT	GGAGCACCAA	GGAGCCCTTT	TCTTGATCA	3660
AGGTGGATCT	GTTGGCACCA	ATGATTATTC	ACGGCATCAA	GACCCAGGGT	GCCCCGTCAGA	3720

AGTTCTCCAG	CCTCTACATC	TCTCAGTTTA	TCATCATGTA	TAGTCTTGAT	GGGAAGAAGT	3780
GGCAGACTTA	TCGAGGAAAT	TCCACTGGAA	CCTTAATGGT	CTTCTTTGGC	AATGTGGATT	3840
CATCTGGGAT	AAAACACAAT	ATTTTAAACC	CTCCAATTAT	TGCTCGATAC	ATCCGTTTGC	3900
ACCCAATCA	TTATAGCATT	CGCAGCACTC	TTCGCATGGA	GTTGATGGGC	TGTGATTTAA	3960
ATAGTTGCAG	CATGCCATTG	GGAATGGAGA	GTAAAGCAAT	ATCAGATGCA	CAGATTACTG	4020
CTTCATCCTA	CTTTACCAAT	ATGTTTGCCA	CCTGGTCTCC	TTCAAAAAGCT	CGACTTCACC	4080
TCCAAGGGAG	GAGTAATGCC	TGGAGACCTC	AGGTGAATAA	TCCAAAAGAG	TGGCTGCAAG	4140
TGGACTTCCA	GAAGACAATG	AAAGTCACAG	GAGTAACTAC	TCAGGGAGTA	AAATCTCTGC	4200
TTACCAGCAT	GTATGTGAAG	GAGTTCCTCA	TCTCCAGCAG	TCAAGATGGC	CATCAGTGGA	4260
CTCTCTTTTT	TCAGAATGGC	AAAGTAAAGG	TTTTTCAGGG	AAATCAAGAC	TCCTTCACAC	4320
CTGTGGTGAA	CTCTCTAGAC	CCACCGTTAC	TGACTCGCTA	CCTTCGAATT	CACCCCCAGA	4380
GTTGGGTGCA	CCAGATTGCC	CTGAGGATGG	AGGTTCGGG	CTGCGAGGCA	CAGGACCTCT	4440
ACTGAGGGTG	GCCACTGCAG	CACCTGCCAC	TGCCGTCACC	TCTCCCTCCT	CAGCTCCAGG	4500
GCAGTGTCCC	TCCCTGGCTT	GCCTTCTACC	TTTGTGCTAA	ATCCTAGCAG	ACACTGCCTT	4560
GAAGCCTCCT	GAATTAACTA	TCATCAGTCC	TGCATTTCTT	TGGTGGGGGG	CCAGGAGGGT	4620
GCATCCAATT	TAACTTAACT	CTTACCTATT	TTCTGCAGCT	GCTCCAGAT	TACTCCTTCC	4680
TTCCAATATA	ACTAGGCAAA	AAGAAGTGAG	GAGAAACCTG	CATGAAAGCA	TTCTTCCCTG	4740
AAAAGTTAGG	CCTCTCAGAG	TCACCACTTC	CTCTGTTGTA	GAAAAACTAT	GTGATGAAAC	4800
TTTGAAAAAG	ATATTTATGA	TGTTGCGGCC	GC			4832

(2) INFORMATION FOR SEQ ID NO:47:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1457 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Met	Gln	Ile	Glu	Leu	Ser	Thr	Cys	Phe	Phe	Leu	Cys	Leu	Leu	Arg	Phe	1	5	10	15
Cys	Phe	Ser	Ala	Thr	Arg	Arg	Tyr	Tyr	Leu	Gly	Ala	Val	Glu	Leu	Ser	20	25	30	
Trp	Asp	Tyr	Met	Gln	Ser	Asp	Leu	Gly	Glu	Leu	Pro	Val	Asp	Ala	Arg	35	40	45	
Phe	Pro	Pro	Arg	Val	Pro	Lys	Ser	Phe	Pro	Phe	Asn	Thr	Ser	Val	Val	50	55	60	
Tyr	Lys	Lys	Thr	Leu	Phe	Val	Glu	Phe	Thr	Asp	His	Leu	Phe	Asn	Ile	65	70	75	80
Ala	Lys	Pro	Arg	Pro	Pro	Trp	Met	Gly	Leu	Leu	Gly	Pro	Thr	Ile	Gln	85	90	95	
Ala	Glu	Val	Tyr	Asp	Thr	Val	Val	Ile	Thr	Leu	Lys	Asn	Met	Ala	Ser	100	105	110	
His	Pro	Val	Ser	Leu	His	Ala	Val	Gly	Val	Ser	Tyr	Trp	Lys	Ala	Ser	115	120	125	
Glu	Gly	Ala	Glu	Tyr	Asp	Asp	Gln	Thr	Ser	Gln	Arg	Glu	Lys	Glu	Asp	130	135	140	
Asp	Lys	Val	Phe	Pro	Gly	Gly	Ser	His	Thr	Tyr	Val	Trp	Gln	Val	Leu	145	150	155	160
Lys	Glu	Asn	Gly	Pro	Met	Ala	Ser	Asp	Pro	Leu	Cys	Leu	Thr	Tyr	Ser	165	170	175	

Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile
 180 185 190
 Gly Ala Leu Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr
 195 200 205
 Gln Thr Leu His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly
 210 215 220
 Lys Ser Trp His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp
 225 230 235 240
 Ala Ala Ser Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr
 245 250 255
 Val Asn Arg Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val
 260 265 270
 Tyr Trp His Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile
 275 280 285
 Phe Leu Glu Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser
 290 295 300
 Leu Glu Ile Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met
 305 310 315 320
 Asp Leu Gly Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His
 325 330 335
 Asp Gly Met Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro
 340 345 350
 Gln Leu Arg Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp
 355 360 365
 Leu Thr Asp Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser
 370 375 380
 Pro Ser Phe Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr
 385 390 395 400
 Trp Val His Tyr Ile Ala Ala Glu Glu Glu Asp Trp Asp Tyr Ala Pro
 405 410 415
 Leu Val Leu Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn
 420 425 430
 Asn Gly Pro Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met
 435 440 445
 Ala Tyr Thr Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu
 450 455 460
 Ser Gly Ile Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu
 465 470 475 480
 Leu Ile Ile Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro
 485 490 495

His Gly Ile Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys
 500 505 510
 Gly Val Lys His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe
 515 520 525
 Lys Tyr Lys Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp
 530 535 540
 Pro Arg Cys Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg
 545 550 555 560
 Asp Leu Ala Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu
 565 570 575
 Ser Val Asp Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val
 580 585 590
 Ile Leu Phe Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu
 595 600 605
 Asn Ile Gln Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp
 610 615 620
 Pro Glu Phe Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val
 625 630 635 640
 Phe Asp Ser Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp
 645 650 655
 Tyr Ile Leu Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe
 660 665 670
 Ser Gly Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr
 675 680 685
 Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro
 690 695 700
 Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly
 705 710 715 720
 Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp
 725 730 735
 Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys
 740 745 750
 Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser Gln Asn Pro Pro Val Leu
 755 760 765
 Lys Arg His Gln Arg Glu Ile Thr Arg Thr Thr Leu Gln Ser Asp Gln
 770 775 780
 Glu Glu Ile Asp Tyr Asp Asp Thr Ile Ser Val Glu Met Lys Lys Glu
 785 790 795 800
 Asp Phe Asp Ile Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe
 805 810 815
 Gln Lys Lys Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp

820	825	830
Asp Tyr Gly Met Ser Ser Ser 835	Pro His Val Leu Arg 840	Asn Arg Ala Gln 845
Ser Gly Ser Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr 850	855	860
Asp Gly Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu His 865	870	875 880
Leu Gly Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp Asn Ile 885	890	895
Met Val Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser Phe Tyr Ser 900	905	910
Ser Leu Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu Pro Arg 915	920	925
Lys Asn Phe Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys Val 930	935	940
Gln His His Met Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys Ala Trp 945	950	955 960
Ala Tyr Phe Ser Asp Val Asp Leu Glu Lys Asp Val His Ser Gly Leu 965	970	975
Ile Gly Pro Leu Leu Val Cys His Thr Asn Thr Leu Asn Pro Ala His 980	985	990
Gly Arg Gln Val Thr Val Gln Glu Phe Ala Leu Phe Phe Thr Ile Phe 995	1000	1005
Asp Glu Thr Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg Asn Cys 1010	1015	1020
Arg Ala Pro Cys Asn Ile Gln Met Glu Asp Pro Thr Phe Lys Glu Asn 1025	1030	1035 1040
Tyr Arg Phe His Ala Ile Asn Gly Tyr Ile Met Asp Thr Leu Pro Gly 1045	1050	1055
Leu Val Met Ala Gln Asp Gln Arg Ile Arg Trp Tyr Leu Leu Ser Met 1060	1065	1070
Gly Ser Asn Glu Asn Ile His Ser Ile His Phe Ser Gly His Val Phe 1075	1080	1085
Thr Val Arg Lys Lys Glu Glu Tyr Lys Met Ala Leu Tyr Asn Leu Tyr 1090	1095	1100
Pro Gly Val Phe Glu Thr Val Glu Met Leu Pro Ser Lys Ala Gly Ile 1105	1110	1115 1120
Trp Arg Val Glu Cys Leu Ile Gly Glu His Leu His Ala Gly Met Ser 1125	1130	1135
Thr Leu Phe Leu Val Tyr Ser Asn Lys Cys Gln Thr Pro Leu Gly Met 1140	1145	1150

Ala Ser Gly His Ile Arg Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr
 1155 1160 1165
 Gly Gln Trp Ala Pro Lys Leu Ala Arg Leu His Tyr Ser Gly Ser Ile
 1170 1175 1180
 Asn Ala Trp Ser Thr Lys Glu Pro Phe Ser Trp Ile Lys Val Asp Leu
 1185 1190 1195 1200
 Leu Ala Pro Met Ile Ile His Gly Ile Lys Thr Gln Gly Ala Arg Gln
 1205 1210 1215
 Lys Phe Ser Ser Leu Tyr Ile Ser Gln Phe Ile Ile Met Tyr Ser Leu
 1220 1225 1230
 Asp Gly Lys Lys Trp Gln Thr Tyr Arg Gly Asn Ser Thr Gly Thr Leu
 1235 1240 1245
 Met Val Phe Phe Gly Asn Val Asp Ser Ser Gly Ile Lys His Asn Ile
 1250 1255 1260
 Phe Asn Pro Pro Ile Ile Ala Arg Tyr Ile Arg Leu His Pro Thr His
 1265 1270 1275 1280
 Tyr Ser Ile Arg Ser Thr Leu Arg Met Glu Leu Met Gly Cys Asp Leu
 1285 1290 1295
 Asn Ser Cys Ser Met Pro Leu Gly Met Glu Ser Lys Ala Ile Ser Asp
 1300 1305 1310
 Ala Gln Ile Thr Ala Ser Ser Tyr Phe Thr Asn Met Phe Ala Thr Trp
 1315 1320 1325
 Ser Pro Ser Lys Ala Arg Leu His Leu Gln Gly Arg Ser Asn Ala Trp
 1330 1335 1340
 Arg Pro Gln Val Asn Asn Pro Lys Glu Trp Leu Gln Val Asp Phe Gln
 1345 1350 1355 1360
 Lys Thr Met Lys Val Thr Gly Val Thr Thr Gln Gly Val Lys Ser Leu
 1365 1370 1375
 Leu Thr Ser Met Tyr Val Lys Glu Phe Leu Ile Ser Ser Ser Gln Asp
 1380 1385 1390
 Gly His Gln Trp Thr Leu Phe Phe Gln Asn Gly Lys Val Lys Val Phe
 1395 1400 1405
 Gln Gly Asn Gln Asp Ser Phe Thr Pro Val Val Asn Ser Leu Asp Pro
 1410 1415 1420
 Pro Leu Leu Thr Arg Tyr Leu Arg Ile His Pro Gln Ser Trp Val His
 1425 1430 1435 1440
 Gln Ile Ala Leu Arg Met Glu Val Leu Gly Cys Glu Ala Gln Asp Leu
 1445 1450 1455
 Tyr

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 100 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

```

Arg Gly Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr
 1              5              10              15
Gly Asp Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu
          20              25              30
Ser Lys Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser Gln Asn Ser Arg
          35              40              45
His Pro Ser Thr Arg Gln Lys Gln Phe Asn Ala Thr Pro Pro Thr Pro
          50              55              60
Pro Thr Pro Pro Val Leu Lys Arg His Gln Arg Glu Ile Thr Arg Thr
          65              70              75              80
Thr Leu Gln Ser Asp Gln Glu Glu Ile Asp Tyr Asp Asp Thr Ile Ser
          85              90              95
Val Glu Met Lys
          100

```

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 300 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

```

AGAGGCATGA CCGCCTTACT GAAGGTTTCT AGTTGTGACA AGAACACTGG TGATTATTAC      60
GAGGACAGTT ATGAAGATAT TTCAGCATAC TTGCTGAGTA AAAACAATGC CATTGAACCA      120
AGAAGCTTCT CCCAGAAATC TAGACACCCT AGCACTAGGC AAAAGCAATT TAATGCCACC      180
CCTCCTACAC CACCAACCCC ACCAGTACTG AAACGCCATC AACGGGAAAT AACTCGTACT      240
ACTCTTCAGT CTGATCAAGA GGAAATTGAC TATGATGATA CCATATCAGT TGAAATGAAG      300

```

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

```

Ser Arg His Pro Ser Thr Arg Gln Lys Gln Phe Asn Ala Thr Pro Pro
 1              5              10              15
Val Leu Lys Arg His Gln Arg Glu Ile Thr Arg
          20              25

```

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 81 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

TCGCGACACC CTAGCACTAG GCAAAAAGCAA TTAAATGCCA CCCCACCACT CCTGAAACGC 60
CATCAACGGG AAATAACGCG T 81

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

ACTACTCTTC AATCTGATCA AGAGGAA 27

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

CGCGCCGCTC GAGTCTACAA TGGCTTTGCC TTTTGCTTTA CTG 43

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

GCGCCCATCG ATTTATTCCT TCCTCCTTAA CCTTCTTGC AAG 43

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 48 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

CGCGCCGCTC GAGCATCCCA ATGGCCCTGT CCTTTCTTT ACTGATGG 48

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

CCATCGATTC AATCCTTCCT CCTTAATCTT TTTTGCAAG 39

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

CCGGATCCTC TACAATGGCT TTGCCTTTTG CTTTACTG

38

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 46 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

CGCGCCGGCG GCCGCTTATT CCTTCCTCCT TAACCTTTCT TGCAAG

46

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

CCGGATCCCA TCCCAATGGC CCTGTCCTTT TCTTTACTGA TGG

43

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 46 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

CGCGCCGGCG GCCGCTCAAT CCTTCCTCCT TAATCTTTT TGCAAG

46

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 90 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

AGCTTGCTGT TTGTGTGCTG CCTCTGAACT CCACACTGAA CAAACTTCAG CCTACTCATG
TCCCTAAAAT GGGCAAACAT TGCAAGCAGC

60

90

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 70 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

AAACAGCAAA CACACAGCCC TCCCTGCGCTG CTGACCTTGG AGCTGGGGCA GAGGTCAGAG
ACCTCTCTGA

60

70

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 82 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

AGCTTCAGAG AGGTCTCTGA CCTCTGCCCC AGCTCCAAGG TCAGCAGGCA GGGAGGGCTG 60
TGTGTTTGCT GTTTGCTGCT TG 82

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 78 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

CAATGTTTGC CCATTTTAGG GACATGAGTA GGCTGAAGTT TGTTCAGTGT GGAATTCAGA 60
GGCAGCACAC AAACAGCA 78

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 90 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

AATTCGCTGT TTGTGTGCTG CCTCTGAAGT CCACACTGAA CAAACTTCAG CCTACTCATG 60
TCCCTAAAAT GGGCAAACAT TGCAAGCAGC 90

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 70 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

AAACAGCAAA CACACAGCCC TCCCTGCCTG CTGACCTTGG AGCTGGGGCA GAGGTCAGAG 60
ACCTCTCTGG 70

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 82 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

AATTCAGAG AGGTCTCTGA CCTCTGCCCC AGCTCCAAGG TCAGCAGGCA GGGAGGGCTG 60
TGTGTTTGCT GTTTGCTGCT TG 82

- (2) INFORMATION FOR SEQ ID NO:68:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 78 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:
CAATGTTTGC CCATTTTAGG GACATGAGTA GGCTGAAGTT TGTTCAGTGT GGAATTCAGA 60
GGCAGCACAC AAACAGCG 78
- (2) INFORMATION FOR SEQ ID NO:69:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 33 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:
CGCGCCGCCC GGGGTAGATC TTGCTACCAG TGG 33
- (2) INFORMATION FOR SEQ ID NO:70:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 40 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:
GCGCCCGCGG CCGCCACTGT CCCAGGTCAG TGGTGGTGCC 40
- (2) INFORMATION FOR SEQ ID NO:71:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 46 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:
CGCGCCGGCG GCCGCTCTAC AATGGCCTTG ACCTTTGCTT TACTGG 46
- (2) INFORMATION FOR SEQ ID NO:72:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 43 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:
GCGCCCATCG ATTCATTCCT TACTTCTTAA ACTTTCTTGC AAG 43
- (2) INFORMATION FOR SEQ ID NO:73:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 46 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: DNA (genomic)
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:
CGCGCCGGCG GCCGCTCTAC AATGGCCTTG ACCTTTGCTT TACTGG 46

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

GCGCCCATCG ATTCATTCCT TACTTCTTAA ACTTCTTGC AAG

43

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

Asn Ser Arg His Pro Ser

1

5

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

Gln Asn Pro Pro Val Leu Lys Arg His Gln Arg Glu Ile Thr

1

5

10

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

AATTCGCGAC ACCCTAGC

18

(2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

CAAAACCCAC CAGTCTTGAA ACGCCATCAA CGGGAAATAA CG

42

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

GCGCTGTGGG ATCGGTTTGT GGTGGTCAGA AC

32

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

TTTGCGGTAG TTGCCCTTTA TTGC

24

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

Arg Thr Leu Gln Ser Asp

1

5

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 15 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

CGTACTCTTC AGTCT

15

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

GCATGAGAAG TCAGACTAG

19